A 1+1 Protection Architecture for Optical Burst Switched Networks

David Griffith, *Member, IEEE,* and SuKyoung Lee

***Abstract*— High-capacity optical backbone networks protect their premium customers' information flows by routing two copies of the customer's data over disjoint paths. This scheme, known as 1+1 protection, provides extremely rapid recovery from network failures. In this paper, we propose an architecture by which 1+1 protection can be extended to Optical Burst Switched (OBS) networks. This architecture is designed by modifying the diversity routing architecture that was originally proposed for non-optical packet networks and recently applied to networks employing the Generalized Multiprotocol Label Switched (GMPLS) architecture. We extend the architecture developed for Just-In-Time (JIT) OBS signaling to support 1+1 protection. We also examine design issues that are caused by a difference in the propagation delays of the two disjoint paths across the OBS network. We show that a sufficiently large difference in the propagation delays can cause performance degradations that may result in an unsatisfactory quality of service (QoS) on the protected connection. We examine the impact of this delay mismatch on restoration performance, probability of burst loss, and jitter. Through analysis and simulations it is discussed how these negative effects can be eliminated.**

***Index Terms*— Optical Burst Switching (OBS), JumpStart JIT, GMPLS, 1+1 Protection, Quality of Service (QoS)**

I. Introduction

MPLS and Generalized MPLS (GMPLS) provide the foundation for a common control plane for many types of transport networks, including optical networks, which allows the development of new kinds of protection architectures [1]. During the past several years proposals have been made in [2], [3], [4] and [5] to use optical restoration mechanisms in MPLS networks. These mechanisms are an extension of Automatic Protection Switching (APS) principles from Synchronous Optical Network (SONET) ring networks to the more general mesh topologies that are being deployed in current-generation Optical Transport Networks (OTNs). APS mechanisms have been discussed extensively in the literature; a summary appears in [6]. Optical protection mechanisms create dedicated backup lightpaths that are disjoint from the working lightpath that normally carries the protected data flow. Optical 1+1 protection reserves resources on two disjoint lightpaths that share common termination nodes and sends the protected data stream over both lightpaths. We call the node closest to the data source, from which the two lightpaths diverge, the protection ingress node, and the node where the two lightpaths merge is the egress node. The egress node acts as a switch and selects data from one lightpath to forward to the destination based on the measured Optical Signal to Noise Ratio (OSNR); the lightpath with the higher OSNR is the one that is effectively connected to the destination. If one lightpath fails, or experiences significant reduction in OSNR, the egress node is able to switch over to the other lightpath nearly instantaneously. This provides protection against both hard and soft failures.

A variation on the 1+1 concept for optical networks has been standardized for data networks using MPLS in [7]. Like 1+1 protection in the optical layer, this approach uses two disjoint paths, but at the MPLS layer. This is an extension of diversity routing, which was originally proposed by Maxemchuck [8]. In terms of restoration time, the 1+1 scheme has a significant advantage [9] over soft protection reservation schemes. However, this method requires the network operator to resolve performance degradation issues due to variations in the delay between the two paths. The challenge is to design a appropriate restoration strategy that synchronizes the two paths.

MPLS 1+1 packet protection can be readily applied to optical burst switched networks, which apply packet switching principles at the optical layer. There have been few works on OBS protection [10],[11]. While wavelength-routed OBS survivability is studied in [10], the work is focused on single link failures. The authors in [11] has applied elements of the MPLS recovery architecture described in [2] to burst switched networks by defining mechanisms for transmitting Failure Indication Signals (FISs) upstream from a point of failure to OBS Path Switch LSRs (OPSLs), which carry out coordinated switching of affected traffic to new paths with their corresponding OBS Path Merge LSRs (OPMLs). Our approach in this paper, which has not been done previously, is to develop an architecture for OBS 1+1 protection based on the MPLS 1+1 protection architecture in [7].

In this paper, we develop analytical models for the most known two burst assembly approaches in literature: threshold-based and timer-based approaches [12], [13] on the basis of which the effect of the delay mismatch between the two paths is investigated. This work is an extension of an earlier analysis by the authors in [14]. Furthermore, we discuss mechanisms for improving the performance of networks using MPLS 1+1 protection. It is important to emphasize that the architecture we are proposing is applicable only to long duration sessions over predefined, pinned routes, and that only premium traffic, comprising a small fraction of the total offered load in an OBS

This research was partially supported by the National Institute of Standards and Technology (NIST), the Advanced R&E Activity (ARDA), the Laboratory for Telecommunications Sciences (LTS) MENTER project, the Defense Advanced Research Projects Agency (DARPA) Fault Tolerant Networks (FTN) program, and the National Communications System (NCS).

D. Griffith is with the National Institute of Standards and Technology (NIST), Gaithersburg, MD 20899 USA (e-mail: david.griffith@nist.gov).

S. Lee is with Sejong University, Seoul, Korea.

network, would be afforded this type of protection, just as 1+1 protection in OTNs is used for only a few customers who are willing to pay for this kind of premium service.

The remainder of this paper is organized as follows. In Section II we describe how 1+1 protection is implemented at the MPLS layer. In Section III we introduce the 1+1 protection architecture for OBS networks. In Section IV we introduce analytical models for the most known two burst assembly approaches: threshold-based and timer-based approaches, and we discuss some of the effects of delay mismatch on the restoration performance of the MPLS 1+1 system and describe some of the resulting design issues. In Section V we examine the effect of delay mismatch on QoS parameters, specifically jitter and the probability of burst loss. In Section VI we provide simulation results that illustrate the impact of delay mismatch on jitter and burst loss rate. We summarize our analysis in Section VII.

II. MPLS 1+1 Protection

The architecture in [7] discusses the basic design principles of MPLS 1+1 protection. 1+1 protection in the transport layer duplicates traffic on two label switched paths that respectively split and merge at ingress and egress Label Switching Routers (LSRs), as shown in Fig. 1. The ingress node is responsible for duplicating packets that are received from the flow source, assigning sequence numbers to them, and sending one copy downstream on each of the two protection Label Switched Paths (LSPs). The egress node is responsible for filtering the two received streams so that only one copy of each packet is forwarded to the flow's destination. This approach is simple to manage and provides fast end-to-end protection. No signaling is required to achieve recovery from link or node failures. If either path is affected by a failure, the egress will continue to receive traffic from the other path. Furthermore, this approach fills a gap that cannot be covered by either Interior Gateway Protocol (IGP)-rerouting which is very slow or MPLS Fast Rerouting (FRR) which does not provide end-to-end protection.

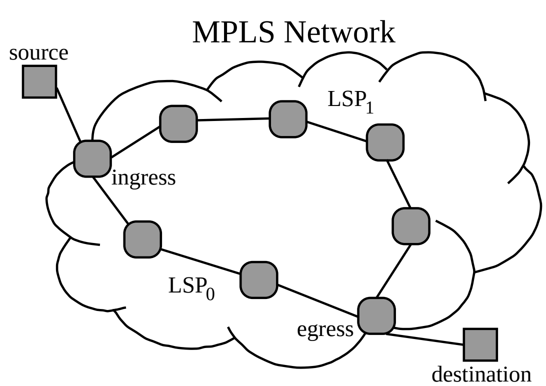

Fig. 1. MPLS 1+1 protection across an MPLS cloud. This provides for continued service in the event of a failure on one of the LSPs.

The MPLS 1+1 protection scheme described in [7] treats both LSPs (e.g. LSP_0 and LSP_1 in Fig. 1) as working paths while traditional MPLS protection [2] designates LSP_0 (or LSP_1) and LSP_1 (or LSP_0) as working and protection paths, respectively. Because the MPLS 1+1 scheme provides a packet level protection service, packets should be buffered to temporally align the two LSPs and compensate for variations in delay between the two paths. The goal is to buffer both paths such that the path that is leading (i.e. whose packets tend to be received by the egress first as defined in [7]) has the same delay as the path that is lagging. In addition, routing algorithms for choosing multiple non-overlapping paths (e.g. the heuristic created by Bhandari [15]) can be modified to choose two LSPs so that the expected propagation delays on the two paths are as close as possible.

The packet selection scheme at the egress is carried out based on the packet sequence number, which is contained in the MPLS shim header, and on the status of a sliding receive window maintained at the egress. Packets are accepted or rejected by the egress LSR based on whether their sequence numbers fall within the range defined by the window at the time of their arrival. If a packet is accepted, the window is adjusted so that its lower limit is one greater than the sequence number of the accepted packet. The operation of the window can be seen in Fig. 2 for the case where the first packet has sequence number 1. In [7], the authors describe several constraints on the range, L, of the window. For instance, L must be large enough so that it is greater than the longest likely burst of lost packets on either LSP, so that the packet sequence numbers do not fall outside the window range and result in all data being lost until the sequence numbers wrap and reenter the window's range from below.

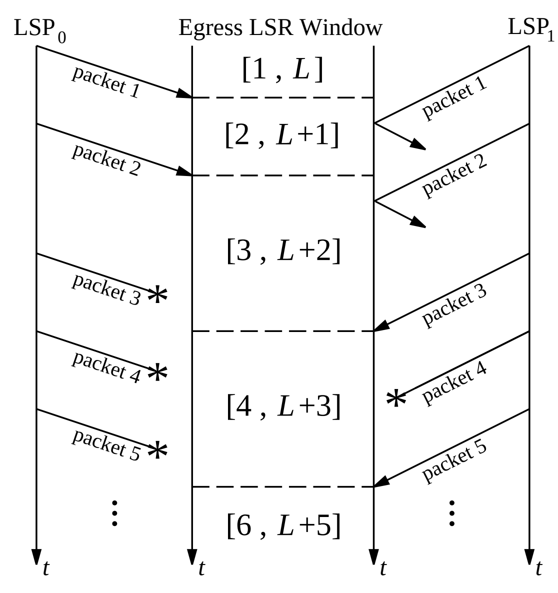

Fig. 2. An illustration of the windowing function used at the egress LSR in a MPLS 1+1 protection system. The window has length L and is adjusted upon each receipt of a packet whose sequence number lies within its range.

III. 1+1 Protection for Optical Burst Switched Networks

A. Background

Optical burst switching provides a mechanism for moving large quantities of bursty data across a transparent optical switching network. OBS can serve as a bridging technology between existing circuit switched transparent optical transport networks and future networks that use pure optical packet switching. In the near term, OBS can be used to increase the efficiency of existing networks, either by using it in the network core to carry out statistical multiplexing of many streams of variable bit rate data, or by using it at the network edge to aggregate multiple traffic streams onto predefined wavelengths in the core network.

Optical burst switching has been described extensively in the literature; [16] and [17] provide good general discussions of how OBS works. We briefly review the essential details here. Bursts are assembled at the edge of the network by accumulating a set of higher layer protocol data units into a large block of data, called an optical burst. The known burst assembly approaches in literature are timer-based [12] and threshold-based [13]. In the latter approach, once the length of a burst being created reaches a threshold value, the burst is generated and transmitted. In the former approach the ingress node generates bursts at regular time intervals; the bursts can be of fixed length or variable length depending on whether the burst that is created when the timer expires is padded into a fixed length burst.

When a burst is created, the edge node creates a control packet and transmits it in advance of the burst, usually on a separate wavelength, to reserve resources for the burst at each switch that it will pass through on the way to its destination. The control packet contains all the information required by each switch to forward the burst, such as the length of the burst and the time offset the between the control packet's arrival at the switch and the arrival of the burst itself. The time delay between the transmission of the control packet and the launching of its associated burst must be large enough to prevent the burst from overtaking the control packet; the minimum value for the time offset is $H\delta$, where H is the number of hops on the burst's path and δ is the mean time required to process the control packet at each switch. Because OBS switches do not convert bursts into electronic form to switch them, the bursts do not experience any queuing delays as they cross the network. The only delays that they encounter are the propagation delays associated with physically traveling along the various fiber links. It is possible for additional delays to occur if multiple bursts arriving at a given switch have to contend for a single output port's resources. Some OBS architectures employ contention resolution mechanisms such as deflection routing or Fiber Delay Lines (FDLs). In this paper we assume that neither approach is being used when we analyze recovery from faults and the effect of delay mismatch on burst loss probabilities; however, we simulate the effect of FDLs on jitter performance in Section V.

B. Signaling

For the OBS 1+1 architecture, we assume that the network uses a signaling architecture, such as the JumpStart Just-In-Time (JIT) described in [18], that allows the creation of semipermanent or permanent connections. The concept of session is first introduced in [18] to provide premium services with persistent route connections. In this paper, the concept is extended to OBS 1+1 protection. We would use OBS 1+1 protection only in the context of a session created by one of the network endpoints; it makes little sense to attempt to use 1+1 protection for individual bursts. The primary reason why the concept of session is introduced in the OBS 1+1 architecture is that high-priority traffic cannot be isolated from lower-priority traffic in OBS networks. Even though there has been some work to try to support QoS, it was offset time that they analyzed to make optical networks capable of QoS support [19]. For a given traffic flow with a requested QoS level, the additional offset time must been chosen large enough to guarantee the resource, but this can cause long end-to-end delay. The proposed OBS 1+1 architecture could reduce this end-to-end delay by establishing a session. We assume that the network is capable of computing disjoint routes and pinning them so that all bursts associated with a session follow the same path. Output ports are reserved

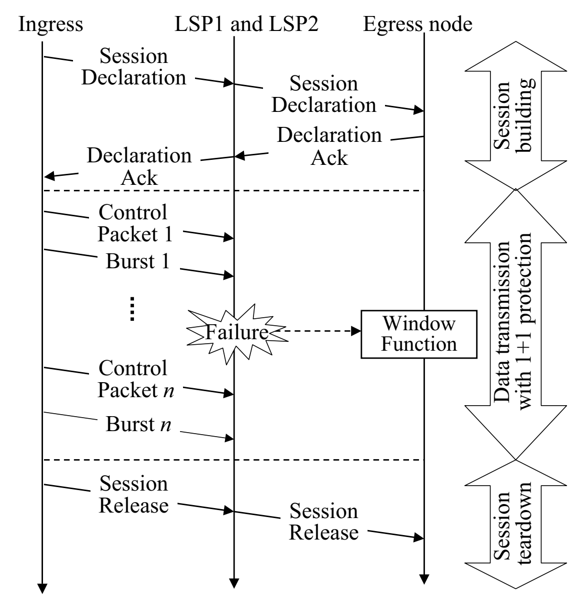

Fig. 3. Session establishment

over the path for a session established to support high-priority bursts. Each control packet supporting the session does not experience any contention for output ports. In the proposed OBS 1+1 architecture, wavelengths are not reserved to support the session; permanently allocating resources is extravagant, effectively resulting in the creation of an optical circuit.

In order to minimize the delay mismatch, the two routes that are chosen to support a protected session should be selected so that the difference between their respective propagation delays is as small as possible. The source node must be aware of the

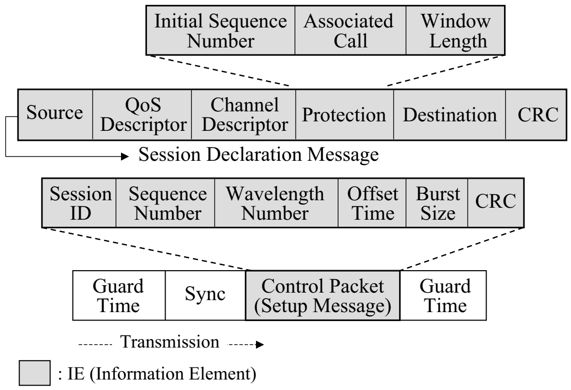

Fig. 4. Control packet including session ID and sequence number

total propagation delay on each route in order to determine what buffering delay must be applied to the leading path.

Once the routes have been computed, the network must insert state information into each of the intermediate switches on both of the paths that will be used to support the protected session. In the JumpStart JIT architecture [18], the Session Declaration message is used to create a long-duration path by installing routing information at each intermediate node. A detailed description of the signaling architecture may be found in [20]. Each field in the message can be either the hardware IE (Information Element) or the software IEs, depending on the network architecture. To support 1+1 OBS protection, the source node must transmit two Session Declaration messages, one on each path as can be seen in Fig. 3. Each Session Declaration message must contain information that will allow the egress node to set up a receive window selective forwarding mechanism. That is, both Session Declaration messages must carry an Explicit Route HardPath IE that contains a list of node addresses corresponding to the intermediate nodes on the route.

The HardPath IEs in the two messages will be the same, except for their Call Reference fields, which allow the call endpoints to differentiate the two paths that support the protected call. To carry the protection information, we define a Protection IE that must be carried in both Session Declaration messages in the block of SoftPath IEs. SoftPath IEs are typically used to convey information between path endpoints, whereas HardPath IEs carry information that is needed by the intermediate OBS switches on a path. In such an environment, the presence of a Protection IE in a Session Declaration message indicates to the egress node that it should create a receive window and other state information to support the protected call. Alternatively, 1+1 protection can be indicated by setting a flag in the QoS IE, which is carried as a HardPath IE in the Session Declaration message.

At a minimum, the Protection IE must consist of <Associated Call, Window Length, Initial Sequence Number> fields as shown in Fig. 4. The latter two fields respectively specify the length of the receive window and the sequence number that will be carried by the first burst. In each Session Declaration message's Protection IE, the ingress sets the Associated Call field to be equal to the Call Reference field in the other Session Declaration message. This allows the egress to create a logical grouping of the two paths into a single protection entity. If the merge point for the two paths is incapable of supporting 1+1 protection, it must send a Failure message back to the ingress node with the appropriate value in the Cause Value field.

The egress node must send a Declaration ACK message back to the ingress for each Session Declaration message that it receives. Once the ingress receives both ACKs, it can begin transmitting bursts. Each burst is duplicated and assigned its own Setup message before being transmitted, which is the control packet explained in the previous section III-A. As shown in Fig. 4, each Setup message that precedes a burst must carry the burst's sequence number in a SoftPath IE so that the egress node can determine whether to forward the burst or to discard it. The intermediate nodes on each path do not need any special functionality to support OBS 1+1 protection; they merely schedule and forward the bursts that they receive. Once both paths are established, the originating node can begin sending data.

If Keepalive messages are used, they must be sent on both paths; they do not need to carry any additional IEs. When the call is over, the ingress node must send Session Release messages on both paths to tear down the state in the intermediate nodes and to remove the receive window and call associations from the egress node.

IV. The Effect of Delay Mismatch on Restoration Performance

Delay mismatch is a measure of the difference in propagation delay and control packet time offset between the two paths used for 1+1 protection, measured in bursts. We assume that the line rates on the two paths are the same and, as we explain later, that control packet offset times are relatively small, so that only the propagation delay difference ΔD (measured in units of time) determines the delay mismatch. We assume without loss of generality that the fixed propagation delays D_0 and D_1 on LSP_0 and LSP_1, respectively, are related as $D_0 < D_1$, so that the delay difference ΔD is given by

$$\Delta D = D_1 - D_0 > 0. \tag{1}$$

Delay mismatch can be minimized by using a constrained K-shortest path algorithm [15] that finds two disjoint paths whose respective propagation delays are as close together as possible. However, this approach cannot completely eliminate the effect of delay mismatch, as the following example demonstrates. Consider a simple network example shown in Fig. 5. Suppose that we want to establish a protected connection from Node 1 to Node 2. If we require that the two paths that will be used to support this connection must be edge-disjoint, then the only two paths that can be used are LSP0 (1-3) and LSP1 (1-2-3). The length of the first path is 384 km, while the length of the second path is 1872 km. This results in a propagation delay difference of $\Delta D = 7.45$ msec. If we assume an average burst size of 15 KBytes (Burst sizes in most OBS schemes are in the KByte range), which was used in [19], with a connection data rate of 622 Mbps, we obtain a delay mismatch of around 38 bursts. As we shall see in the following discussion, the effects associated with such a mismatch can be significant. Even at lower data rates, the delay mismatch between the two paths can be unacceptably large, as the graph in Fig. 6 shows. The graph shows contour lines associated with the delay mismatch given by the function $\Delta D(b/s)$, where b is the mean data rate and s is the mean burst size. As the figure shows, at a data rate of around 100 Mbps, a delay mismatch of less than one burst length is possible only if the session uses an average burst size of at least 100 KBytes. Higher data rates require even larger bursts in order to keep the delay mismatch low. This may not be possible in general because of the additional delay variation that it introduces. The behavior of MPLS 1+1 protection

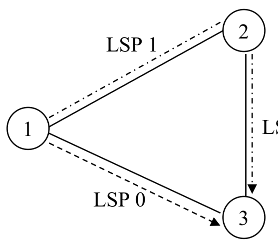

Fig. 5. An example network with two paths

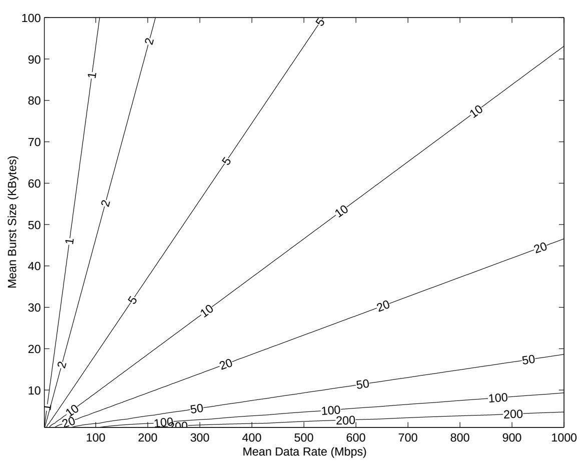

Fig. 6. Delay mismatch as a function of burst size and session data rate when $\Delta D = 7.45$ msec.

during LSP outage events was discussed in [7], which noted that there will be a delay in new packet arrivals if a failure occurs on the leading (i.e. less delayed) LSP. In this section we quantify this behavior for OBS networks where burst assembly is an essential part at the ingress nodes.

A. Assumptions

If H_0 and H_1 are the respective number of hops on the two paths which we will designate as LSP_0 and LSP_1, the bursts on LSP_0 and LSP_1 must initially lag their Setup messages by at least $H_0\delta$ and $H_1\delta$ seconds, respectively.

The egress node maintains a receiver window of length L, as discussed in Section II. When a control packet arrives at this node, its sequence number is checked and compared to the current range of the receive window. If the sequence number lies within the window's range, the egress node will forward the burst associated with that control packet. Otherwise, the burst will be dropped. If a burst is to be accepted, the receive window will advance, so that if sequence number n is accepted, the new window range will be $[n+1, n+L]$. As with 1+1 MPLS protection, the sequence numbers occupy a fixed length field of B bits, so that they take a maximum value of $2^B - 1$ and then wrap around to zero.

Because bursts are dropped when contention occurs, we can assign a loss probability of p_0 to bursts on LSP_0 and a loss probability of p_1 to bursts on LSP_1. These probabilities represent the total probability of loss due to contention on each path. We assume that the probability that any burst is lost is independent of the probability that any other burst is lost, and that the network is in steady state so that the loss probabilities are not time varying. The independence assumption has been employed in previous work on OBS networks and has been shown to be justified. For example, in [21], which characterizes blocking probabilities in OBS networks, the authors used simulations to validate their assumption that burst blocking events on any link on any path are independent from blocking events on any other link. Furthermore, if session establishment is used to commit streams of bursts to fixed routes, regardless of whether they use 1+1 protection, lost bursts in one part of the network will not induce a migration of traffic to other parts of the network, as would occur in a packet network using adaptive routing.

Because we are using disjoint paths, the propagation delays experienced by the two copies of the burst will be different, even if there is the same number of hops on each path. If the number of hops on LSP_0 is different from the number of hops on LSP_1, then the difference in the arrival times of the two copies of the burst will be $|\Delta D + (H_1 - H_0)\delta|$. In some networks, the ingress node may apply an additional delay to premium traffic in order to isolate it from lower priority traffic, as described in [17]. If only two traffic classes are defined in the network, assigning an additional delay to premium traffic that is equal to the maximum length of a low-priority burst will produce nearly perfect isolation of the two traffic classes. It is reasonable to assume that the same isolation delay will be applied to both paths that are used to support 1+1 protection, and that as a result the isolation delay does not contribute to the delay mismatch between the two paths. If $H_1 < H_0$, meaning that LSP_1 has fewer hops than LSP_0 but the distance a burst travels on LSP_1 is greater than that on LSP_0, the delay mismatch will be less than ΔD. If the average time required to process and forward a Setup message is much less than the average link propagation delay, then we can consider only the link propagation delays when computing the delay mismatch, which we approximate by ΔD.

In order that our analytic model should capture the behavior of the two burst assembly approaches explained in Section III, we adopt the following two system models. The first model is deterministic and characterizes a timer-based approach with padding, or a threshold-based approach where the source is transmitting at a constant (peak) bit rate. The second model uses Poisson streams to model threshold-based burst assembly at an ingress node fed by a source or set of sources that transmit variable bit rate data (variable size packets or packets with random interarrival times).

For both models, we postulate a network in which FDLs are not used in core OBS nodes. Nor are bursts partially dropped if output port contention occurs; contention results in the complete loss of a burst that tries to make use of a busy output port. For each burst, access to a desired output port is reserved in advance for each burst by a control packet that is processed at each intermediate OBS switch; a successful reservation results in the switch fabric's being configured so that the associated burst passes through the switch without undergoing any detection and retransmission (O/E/O conversion). If its associated control packet is lost, a burst will be dropped; an OBS switch will not forward a burst for which it does not have a reservation. A control packet can be lost if it encounters an OBS switch processor module that is busy processing another control packet and that does not have sufficient buffer capacity to enqueue the new arrival.

Under these assumptions, bursts belonging to different sessions that use a common output port on a given OBS switch may interact by causing losses due to contention, but the surviving bursts do not affect each other's propagation delays. Burst flows are therefore cut through the OBS switch network and experience no buffering delays or changes in the interarrival spacing of their constituent bursts as they traverse the network. The egress edge node for a given session will see a stream of bursts whose interarrival times are the same as when the bursts were injected into the OBS network by the ingress edge node that assembled them. Bursts that are lost due to contention will leave empty spaces in the stream that is seen by the egress, but the temporal arrangement of the surviving bursts is not affected by any burst losses. For these reasons, a deterministic or Poisson burst stream that enters the OBS switching network will retain its statistical properties and can be modeled at the egress as, respectively, a deterministic or Poisson stream with holes corresponding to lost bursts.

An example of burst stream behavior is shown in Fig. 7. A sequence of bursts is plotted on the time axis at both the ingress and egress points for the stream to which the bursts belong. The bursts that arrive at the egress are shifted forward in time by a fixed propagation delay, D, so that burst m, launched into the OBS network at time $t = a_m$, arrives at the egress at time $t = a_m + D$. In the figure, the second burst in the stream, shown as a dashed arrow, has been lost somewhere in the OBS network. The temporal spacing of the surviving bursts that arrive at the egress is the same as it was at the ingress. In the figure, the interarrival time between bursts $m-1$ and m is X_m; even though burst 2 was lost, the time delay between the arrivals of burst 1 and burst 3 is still $X_2 + X_3$. We will make extensive use of this property in our later analysis.

B. Restoration for Timer-Based Burst Assembly with Padding

In the deterministic model, we assume that the burst transmission rate λ is the same on the two LSPs. Bursts are uniformly spaced along the time axis on each LSP, with an inter-burst spacing of $1/\lambda$ seconds. The arrivals of bursts on

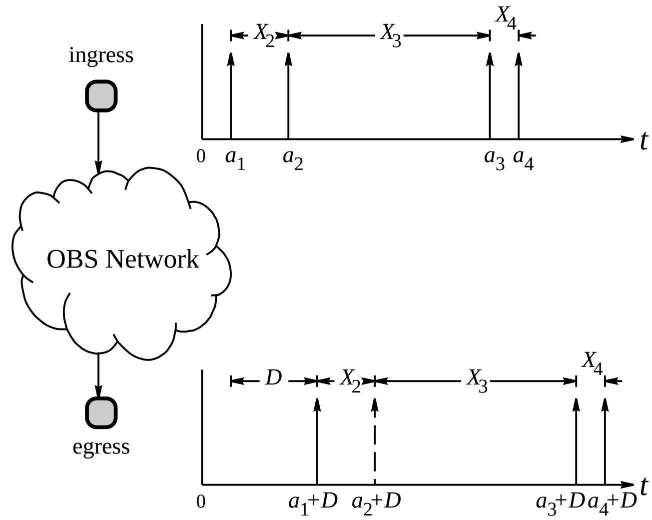

Fig. 7. Burst arrivals plotted versus time at ingress and egress points of an OBS network.

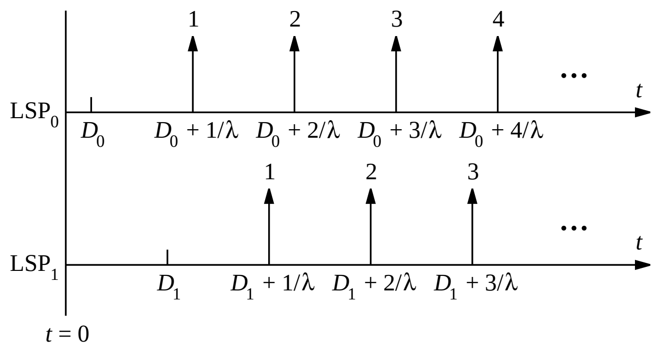

Fig. 8. Deterministic OBS 1+1 system with equal line rates on the two LSPs. Sequence numbers start at 1 in this example.

the two LSPs are shown in Fig. 8. The arrival time of the n^{th} burst on LSP_i is $D_i + n/\lambda$. From the figure, we see that D_1 is related to D_0 and λ as

$$D_0 + \frac{k}{\lambda} < D_1 < D_0 + \frac{k+1}{\lambda}, \tag{2}$$

where k is a non-negative integer given by $k = \lfloor \lambda \Delta D \rfloor$. Under these assumptions, the n^{th} burst arrival on LSP_1 at the egress occurs between the arrival times for the $(k+n)^{\text{th}}$ and $(k+n+1)^{\text{th}}$ bursts on LSP_0.

If LSP_1 fails, there is no effect on the burst stream at the egress, because every burst passed downstream by the egress node is pulled from LSP_0 in this model. If a failure occurs on LSP_0 at time $t = T$, there will be a delay between the last burst received on LSP_0 before the failure event (call this burst n, where $n = \lfloor \lambda (T - D_0) \rfloor$) and the first burst received on LSP_1 and forwarded downstream by the egress. Assuming that no bursts are lost on LSP_1 after the failure event, the first burst received by the egress LSR from LSP_1 after the failure of LSP_0 is burst $n-k$. This burst is discarded by the egress LSR because the egress sequence number window covers the range $[n+1, n+L]$ after the receipt of burst n from LSP_0.

The egress continues to discard bursts until it receives burst $n+1$ from LSP_1; the egress will discard a total of $k+1$ bursts from LSP_1 between the failure of LSP_0 and the resumption of traffic using bursts from LSP_1. The time between the arrival of burst n on LSP_0 and the arrival of burst $n-k$ on LSP_1 is $\Delta D - k/\lambda$. The time gap between the arrival of burst $n-k$ on LSP_1 and the arrival of burst $n+1$ on LSP_1 is $(k+1)/\lambda$. Thus the total time lag between the arrival of the last burst on LSP_0 and the arrival of a burst on LSP_1 that is forwarded downstream is $\Delta D + 1/\lambda$.

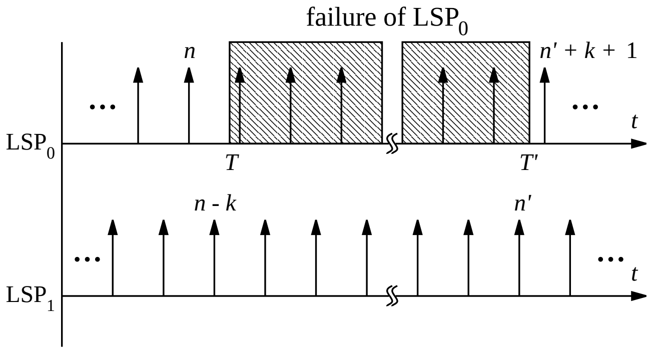

Fig. 9. Illustration of failure event on LSP_0, beginning at time $t = T$ and ending at time $t = T'$.

When service on LSP_0 is restored, a burst will arrive at the egress from LSP_0 after the receipt of burst n' on LSP_1. From (2), the first burst received on LSP_0 after restoration of service will be burst $n' + k + 1$. When this burst is received, the window range is $[n'+1, n'+L]$. There are two possible outcomes. If $k \leq L-1$, then the burst from LSP_0 is accepted and the window advances so that it covers the range $[n'+k+2, n'+k+L+1]$. Subsequently received bursts from LSP_1 will be dropped by the egress; thus, k bursts have been lost. If $k \geq L$ (or, equivalently, if $\Delta D > L/\lambda$), then burst $n'+k+1$ from LSP_0 is rejected by the egress and burst $n'+1$ is received from LSP_1 and passed downstream. The egress will continue to forward bursts from LSP_1 and reject bursts from LSP_0, even though the bursts from LSP_1 are arriving later than their copies that were forwarded over LSP_0. This has serious consequences in the event that a failure occurs on LSP_1, for in that case no bursts will be received from LSP_1 while the egress continues to reject bursts from LSP_0 because their sequence numbers lie outside the range of the egress' receive window. If the transmission continues for a sufficiently long period of time, the situation will be resolved by the wrapping of the sequence numbers so that bursts are accepted from LSP_0, but this may involve the loss of a considerable amount of data. This phenomenon is described in Appendix IV of [7], using results obtained independently from ours, and which also recommends setting L equal to the maximum number of packets by which the lagging LSP can fall behind the leading LSP.

C. Restoration for Threshold-Based Burst Assembly

In this subsection, we consider the case where we have burst streams with Poisson interarrival times since each burst is sent into a network once its length reaches to a threshold value under a threshold-based burst assembly approach. Here the n^{th} burst arrives on LSP_0 at time

$$a_n = D_0 + \sum_{k=1}^{n} X_k, \tag{3}$$

as shown in Fig. 10. The random variables $\{X_i\}_{i=1}^{\infty}$ form a sequence of independent, identically distributed exponential random variables with mean $1/\lambda$. The corresponding arrival time of burst n on LSP_1 is given by $b_n = a_n + \Delta D$.

Suppose that a failure occurs on LSP_0 at some time $t = T$. Let a_m be the time of arrival of the last successfully received burst on LSP_0 before the failure. As was the case with the deterministic model, the egress will discard bursts arriving on LSP_1 until burst $m+1$ arrives. Since burst m arrives on LSP_1 at time $b_m = a_m + \Delta D$, the mean time lapse between $t = a_m$ and $t = b_{m+1}$ is $\Delta D + 1/\lambda$. This result can be generalized to any arrival process whose interarrival times can be represented as a sequence of random variables $\{X_i\}$ whose elements are independent and have the same distribution as some random variable X. It follows directly that the mean time lapse between the last arrival from LSP_0 and the first arrival from LSP_1 that gets forwarded is $\Delta D + \mathsf{E}\{X\}$.

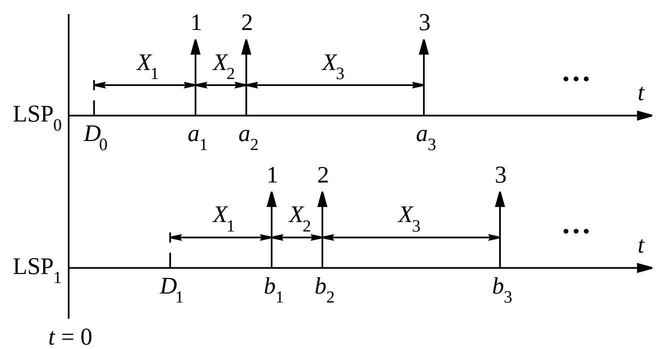

Fig. 10. Poisson streams on two LSPs. The stream on LSP_1 is identical to the stream on LSP_0 but is delayed with respect to it by ΔD units of time.

We can also determine the number of bursts that are dropped after the failure is restored at time $t = T'$. Let $b_{m'}$ be the time of the last arrival on LSP_1 before LSP_0 is restored. The time when burst m' would have arrived from a properly functioning LSP_0 is $a_{m'} = b_{m'} - \Delta D$. We are interested in the average number of arrivals that would have occurred on LSP_0 in the interval $[a_{m'}, T']$, since the first arrival on LSP_0 after time $t = T'$ will be accepted by the egress, and all subsequent arrivals on LSP_1 will be rejected.

Since the stream on LSP_1 is a delayed copy of the stream on LSP_0, it follows that there are no arrivals from LSP_0 in the subinterval $[a_{m'}, a_{m'} + (T' - b_{m'})]$. The remaining subinterval $[a_{m'} + (T' - b_{m'}), T']$ has length ΔD; since the stream is Poisson, the average number of arrivals that occur in this interval is $\lambda \Delta D$. Thus the expected sequence number of the first burst to appear on LSP_0 after it is restored will be $m' + 1 + \lfloor \lambda \Delta D \rfloor$.

After burst m' is received on LSP_1, the egress receive window will cover the range $[m'+1, m'+L]$. If the sequence number of the first received burst on LSP_0 is larger than $m'+L$, then the burst will be rejected by the egress, which will continue to receive bursts from LSP_1. This event will occur if the number of arrivals on LSP_0 in the interval $[T'-\Delta D, T']$ is greater than $L-1$. The probability of this event is

$$\mathsf{Pr}\left\{\text{overshoot}\right\} = \sum_{k=L}^{\infty} \frac{[\lambda \Delta D]^k}{k!} e^{-\lambda \Delta D} = \frac{\gamma(L, \lambda \Delta D)}{\Gamma(L)}, \quad (4)$$

where $\gamma(L, \lambda\Delta D) = \int_0^{\lambda\Delta D} u^{L-1} e^{-u} du$ is the lower incomplete gamma function.

In Fig. 11, for several values of delay mismatch, we plot the probability that the burst stream on LSP_0 will overshoot the upper limit of the receive window versus the receive window length, L. From the figure, we see that a relatively small window size yields low window overflow probabilities when the delay mismatch between the two LSPs is small. Conversely, a large delay mismatch results in an overflow probability close to unity over a large range of window sizes. Also the rate of decrease of the overflow probability is greater when a delay mismatch is smaller. Note that a larger window size is required here to obtain a reasonably low probability that overshoot will not occur, versus the situation a the deterministic model, where setting $L > \lambda\Delta D$ is sufficient to prevent overshoot. Whether the window size is driven by this criterion depends on how severe other effects (such as burst losses) are.

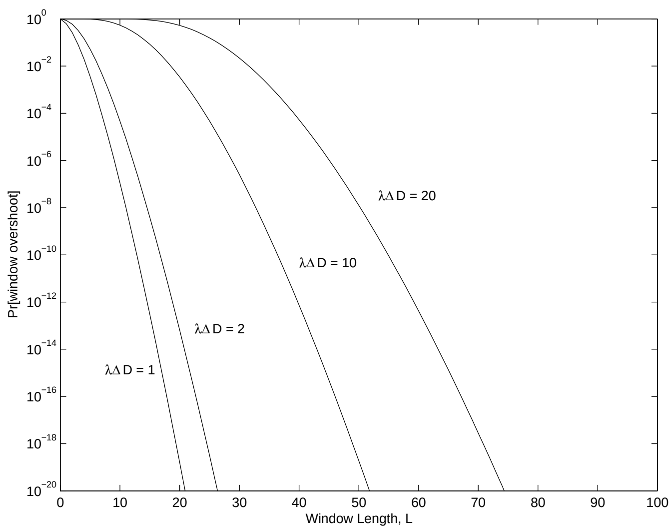

Fig. 11. Probability of window overshoot versus window length, L, for various values of delay mismatch, $\lambda\Delta D$.

V. The Effect of Delay Mismatch on QoS

A. The Effect on Jitter

Instantaneous packet jitter is defined in [22] using the difference in the delay times of two sequentially received bursts as measured by the receiving node. If S_i is the time at which burst i was sent and R_i is the time when burst i is received, then the delay difference between bursts i and j is

$$\begin{aligned} D_{i,j} &= (R_j - R_i) - (S_j - S_i) \\ &= (R_j - S_j) - (R_i - S_i). \end{aligned} \quad (5)$$

Jitter is measured using an adaptive process in which the measured delay difference between sequentially received bursts is the forcing function. The adaptation function for the jitter measurement, given in [22], is

$$J_n = J_{n-1} + \frac{|D_{n-1,n}| - J_{n-1}}{16}. \quad (6)$$

If the jitter process $\{J_n\}$ and the delay difference process $\{D_{n-1,n}\}$ are stationary, then the expected jitter can be found using the expected delay difference between sequentially received bursts at the egress node. In the deterministic model, the delay difference between bursts received on a given LSP is zero. One obtains non-zero jitter measurements in the deterministic system due to random burst losses on each LSP or due to LSP failures, in which case the changes in jitter are transient in nature.

For both models, the expected delay variation between bursts $n-1$ and n is

$$\begin{aligned} \mathsf{E}\left\{|D_{n-1,n}|\right\} &= \mathsf{E}\left\{|D_{n-1,n}|\,|\text{diff}\right\} \mathsf{Pr}\left\{\text{diff}\right\} \\ &\quad + \mathsf{E}\left\{|D_{n-1,n}|\,|\text{same}\right\} \mathsf{Pr}\left\{\text{same}\right\}. \end{aligned} \quad (7)$$

Here, $\mathsf{Pr}\{\text{diff}\}$ is the probability that burst $n-1$ and burst n arrived at the egress from different LSPs. Likewise, $\mathsf{Pr}\{\text{same}\}$ is the probability that the two bursts arrived from the same LSP. We derive both probabilities in the appendix, and find that for the deterministic model,

$$\mathsf{E}\left\{J\right\} = \frac{2(1-p_1)(1-p_0)}{1-p_0 p_1} \Delta D p_0^{\lceil \lambda \Delta D \rceil}. \quad (8)$$

We next compute the average jitter for the Poisson model. For the case where bursts are arriving from different LSPs, we have two bursts with sequence numbers i and j arriving at the egress, and having respective transit times $D_0 + X_i$ and $D_1 + X_j$. The mean absolute difference in their delays is

$$\mathsf{E}\left\{|D_{n-1,n}|\,|\text{diff}\right\} = \mathsf{E}\left\{|X_j - X_i + \Delta D|\right\}, \quad (9)$$

which we can compute as

$$\begin{aligned} &\mathsf{E}\left\{|X_j - X_i + \Delta D|\right\} \\ &\quad = \int_0^\infty \int_0^{x_j + \Delta D} (x_j - x_i + \Delta D) f_{X_i, X_j}(x_i, x_j)\, dx_i dx_j \\ &\qquad + \int_0^\infty \int_{x_j + \Delta D}^\infty (x_i - x_j - \Delta D) f_{X_i, X_j}(x_i, x_j)\, dx_i dx_j \\ &\quad = \Delta D + \frac{e^{-\lambda \Delta D}}{\lambda}, \end{aligned} \quad (10)$$

where we have used the fact that X_i and X_j are independent, exponentially distributed random variables with mean $1/\lambda$. If the two bursts i and j are from the same LSP, then their mean absolute delay difference is

$$\mathsf{E}\left\{|D_{n-1,n}|\,|\text{same}\right\} = \mathsf{E}\left\{|X_j - X_i|\right\} = \frac{1}{\lambda}. \quad (11)$$

In the appendix, we carry out the computations to produce

$\Pr\{\text{same}\}$ and $\Pr\{\text{diff}\}$. They are respectively given by

$$\begin{aligned}
&\Pr\{\text{same}\} \\
&= (1-p_0)\left(1-e^{-\lambda(1-p_0)\Delta D}\right) \\
&\quad + \frac{p_0^2(1-p_1)^2}{1-p_0p_1} e^{-\lambda(1-p_0)\Delta D} \\
&\quad + (1-p_0)^2 \int_{\lambda\Delta D}^{\infty}\int_{\lambda\Delta D/p_1}^{\infty} e^{-(t+u)} I_0(2\sqrt{p_0p_1tu})\, dtdu,
\end{aligned} \tag{12}$$

and

$$\begin{aligned}
&\Pr\{\text{diff}\} \\
&= \frac{(1-p_0)(1-p_1)p_0e^{-\lambda(1-p_0)\Delta D}}{1-p_0p_1} \\
&\quad + (1-p_0)(1-p_1)p_0e^{-\lambda(1-p_1^{-1})\Delta D} \\
&\quad \times \int_{\lambda\Delta D/p_1}^{\infty}\int_{\lambda\Delta D}^{\infty} e^{-(t+u)} I_0(2\sqrt{p_0p_1tu})\, dtdu.
\end{aligned} \tag{13}$$

B. The Effect on Burst Loss Probability

By sending duplicate copies of each burst over disjoint paths, OBS 1+1 protection allows for a considerable reduction in the burst loss rate, in addition to providing a method for rapidly recovering from failure on either of the LSPs. In situations where the delay and transmission rates of the two paths are closely matched, the burst loss rate can trivially be shown to be p_0p_1, where p_0 and p_1 are the loss rates on LSP_0 and LSP_1, respectively, on the assumption in section IV-A. However, if there is a significant difference in the propagation delays associated with the two LSPs, then the probability of burst loss can actually be higher, due to the existence of an additional burst loss mechanism that we analyze here. A burst will be lost if each copy of it is dropped in transit. Bursts can also be lost if one copy is dropped in transit and the other copy is rejected by the egress LSR because its sequence number lies outside the range defined by the sliding window. This will happen if additional bursts arrive on the LSP that dropped the burst, advancing the window, before the undropped copy arrives from the other LSP.

For both models that we introduced in Section IV with independent burst losses on each LSP, we find by conditioning on the LSP burst loss events that

$$\begin{aligned}
\Pr\{\text{loss}\} &= p_0p_1 + p_0(1-p_1)\Pr\left\{\text{loss}|\mathcal{L}_0 \cap \overline{\mathcal{L}_1}\right\} \\
&\quad + p_1(1-p_0)\Pr\left\{\text{loss}|\overline{\mathcal{L}_0} \cap \mathcal{L}_1\right\}
\end{aligned} \tag{14}$$

where the events $\mathcal{L}_0$ and $\mathcal{L}_1$ occur when a burst is lost on LSP_0 and LSP_1, respectively. Now, $\Pr\left\{\text{loss}|\overline{\mathcal{L}_0} \cap \mathcal{L}_1\right\} = 0$ because any burst that arrives from LSP_0 in this model will not be discarded; it always appears before its counterpart arriving from LSP_1.

First we consider the deterministic case. If we let the burst of interest have sequence number n, we see that $\Pr\left\{\text{loss}|\mathcal{L}_0 \cap \overline{\mathcal{L}_1}\right\}$ is the probability that there is at least one successfully received burst from LSP_0 before time $t = D_1 + n/\lambda$, the time when the copy of burst n arrives from LSP_1. From (2), we see that the n^{th} arrival on LSP_1 occurs between the $(k+n)^{\text{th}}$ and $(k+n+1)^{\text{th}}$ arrivals on LSP_0:

$$D_0 + \frac{k+n}{\lambda} < D_1 + \frac{n}{\lambda} < D_0 + \frac{k+n+1}{\lambda}. \tag{15}$$

The only way for the window not to advance and make sequence number n be out of range is for the bursts with sequence numbers $n+1$, $n+2$, $\ldots$, $n+k$ to be dropped by LSP_0. This will occur with probability $\Pr\left\{\text{loss}|\mathcal{L}_0 \cap \overline{\mathcal{L}_1}\right\} = 1-p_0^k$. Thus the probability of burst loss at the egress LSR is

$$\Pr\{\text{loss}\} = p_0 - (1-p_1)p_0^{k+1}, \tag{16}$$

where $k = \lfloor \lambda\Delta D \rfloor$ is the relative offset in bursts of the streams on the two LSPs.

When we have exponential interarrival times, we find that event $\mathcal{L}_0 \cap \overline{\mathcal{L}_1}$ occurs when there is at least one successfully received burst on LSP_0 in the interval $[a_m, b_m]$, given the m^{th} burst was lost on LSP_0. Since $b_m - a_m = \Delta D$, it follows that

$$\begin{aligned}
\Pr\left\{\text{loss}|\mathcal{L}_0 \cap \overline{\mathcal{L}_1}\right\} &= \sum_{k=1}^{\infty}(1-p_1^k)\frac{[\lambda\Delta D]^k}{k!}e^{-\lambda\Delta D} \\
&= 1 - e^{-\lambda(1-p_1)\Delta D}.
\end{aligned}$$

Inserting this result into (14), we obtain

$$\Pr\{\text{loss}\} = p_0[1-(1-p_1)e^{-\lambda(1-p_0)\Delta D}], \tag{17}$$

since $\Pr\left\{\text{loss}|\overline{\mathcal{L}_0} \cap \mathcal{L}_1\right\} = 0$.

For both models, we see that when the delay difference between LSPs is less than one burst interval, the loss probability is just p^2, the probability that both copies of a burst are lost. Once the delay difference exceeds one burst period, the loss behavior of the system in both models approaches that of the leading LSP, which is LSP_0 in this case. In addition, it is clear that the two LSPs should be routed, if possible, so that the LSP with the higher loss rate is also the one with the longer average burst delay.

VI. Results and Discussion

The results from analysis and simulation are presented in this section. We simulated a OBS 1+1 system in which the burst streams arriving at the egress LSR were copies of a single Poisson process where the average burst size is 15 KBytes mentioned in section III. In our simulation system, the sliding window was implemented at the egress with length 100. The burst loss probability was $p = 10^{-6}$ on both LSPs in all simulated cases. In each case, there was little significant deviation from the jitter associated with a single Poisson stream. The jitter was computed over runs of 5000 bursts each, and the curves are ensemble averages taken over 100 runs.

A. Delay Jitter Performance

A plot of the normalized expected jitter $\lambda\mathsf{E}\{J\}$, which is measured in burst intervals, that is introduced into the deterministic arrivals system by burst error is given in Fig. 12 for the case where $p_0 = p_1 = p$. The jitter is plotted versus burst loss probability for three values of normalized delay offset, $\lambda\Delta D$. The peak jitter occurs when $p = 0.4$, and does

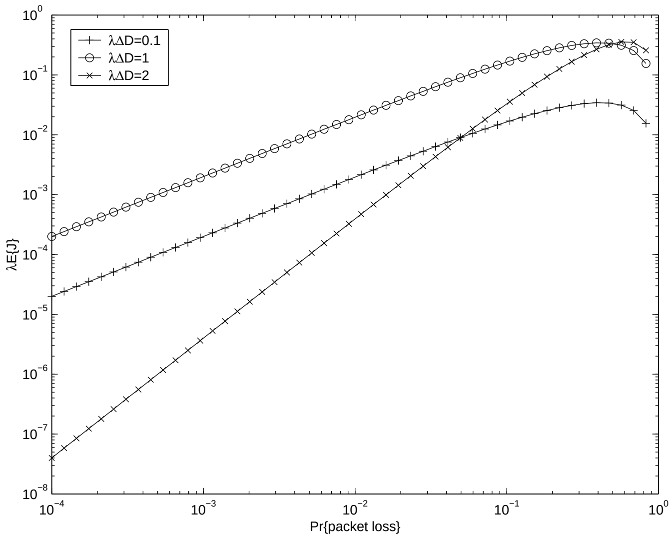

Fig. 12. Normalized jitter in a OBS 1+1 system with deterministic burst arrival times, for different values of $\lambda\Delta D$.

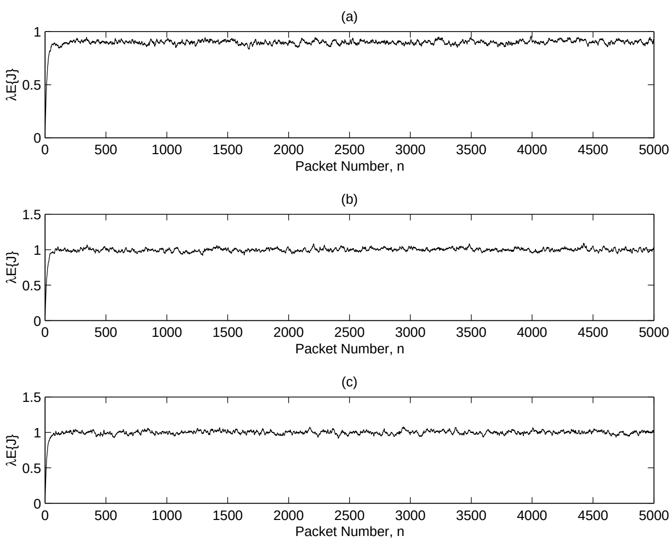

Fig. 13. Plots of average normalized jitter in a OBS 1+1 system with random burst loss and exponential burst interarrival times. **(a)**: $\lambda\Delta D = 1$, **(b)**: $\lambda\Delta D = 10$ **(c)**: $\lambda\Delta D = 100$

not exceed a single burst period. For larger values of delay mismatch, the impact on jitter performance decreases, due to the fact that consecutive forwarded bursts' being taken from different LSP's becomes less likely. Thus it appears that a system using fixed size bursts transmitted at regular intervals will not experience a significant amount of jitter. This assumes, however, that no deflection routing or fiber delay lines are used to resolve contention issues. As we discuss below, the introduction of such mechanisms can significantly increase the level of jitter in the protected stream.

Fig. 13 plots the simulated jitter performance of the Poisson model for three values of $\lambda\Delta D$. This figure as well as the analytical result in Fig. 12 verifies Eq. 8. An interesting result is the apparent reduction in the normalized jitter from unity in the case where $\lambda\Delta D = 0.01$. The deviation is small, and is due to events where a burst arrives from LSP_0 when its

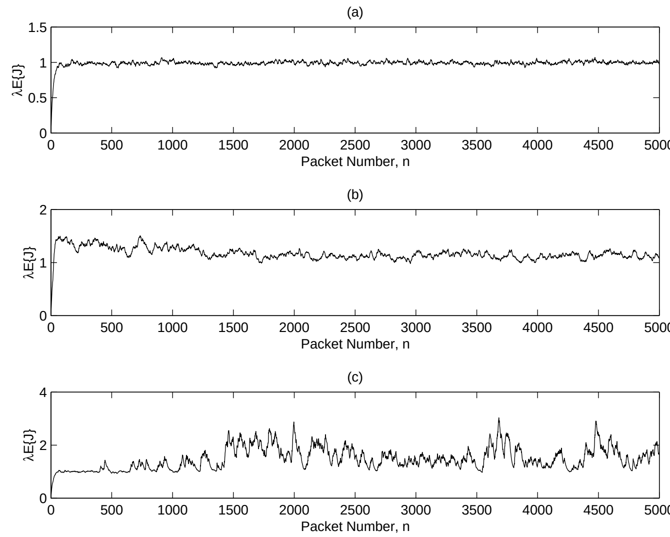

Fig. 14. Plots of average normalized jitter in a OBS 1+1 system with random burst loss and independent Poisson streams on the two LSPs. **(a)**: $\lambda\Delta D = 1$, **(b)**: $\lambda\Delta D = 10$ **(c)**: $\lambda\Delta D = 100$

predecessor arrived from LSP_1. This event becomes far less likely when the delay difference between the two LSPs is on the order of $\lambda\Delta D/10$, as can be seen from Fig. 13(b).

We arrive at a different result if we introduce random delays into the model in addition to the fixed propagation delays. If we suppose that there is a single bottleneck FDL node on each LSP, and model this node using a M/M/1 queue, we introduce an exponentially distributed waiting time that is added to the propagation time of each burst. If we consider our fixed size burst source in this context, it follows that the duplicated burst streams seen by the egress will appear as independent Poisson streams. The analysis of the jitter performance of the system is rather involved, but we can readily simulate its behavior. We show the impact of delay mismatch on jitter in OBS 1+1 systems with random line delays and fixed burst size in Fig. 14. We used the same simulation parameters that we used to generate the results in Fig. 13. In Fig. 14(a), the delay difference is on the order of a single burst period, and the average jitter is very close to the mean burst interarrival time, as we would expect. Increasing the delay difference to 10 average burst intervals produced a greater initial overshoot, slower convergence of $\{\lambda J_n\}$, and a long-term offset of approximately 10% from the average burst interarrival time, as shown in Fig. 14(b). In Fig. 14(c), the delay difference is 100 burst intervals. Because of the long delay and the low burst loss probability, we do not see the impact of the delay difference until the 500^{th} burst. After this point, the jitter curve becomes very noisy, approaching three times the mean burst interarrival time at places. Reducing the size of the sliding window helps only in cases where the delay difference is very large. For $\lambda\Delta D = 200$, we found that if $L < 40$ the jitter curve is well behaved, but setting $L = 4$ for the case where $\lambda\Delta D = 100$ did not eliminate the noise. Thus buffering at the ingress seems to be the best solution.

B. Burst Loss Performance

We now examine the burst loss performance. The average number of bursts that need to be buffered is $\lambda\Delta D$ regardless of the burst assembly mechanism that the ingress node is using. For the timer-based burst assembly model, exactly $B = \beta\lambda\Delta D$ bytes are required for the buffer, where β is the number of bytes in a burst. For the case where threshold-based burst assembly is used, there will be an average of B bytes of data in the buffer at any given time, although the amount of data in the buffer can vary. Therefore, more than B bytes should be made available to prevent bursts' being dropped due to lack of buffer space. Given a buffer with size B, the probability that a burst will be dropped is the probability that more than B/β bursts are generated in ΔD seconds. If we assume that B is an integer multiple of β, this is just

$$p_{\text{overflow}} = \sum_{k=B/\beta+1}^{\infty} \frac{[\lambda\Delta D]^k}{k!} e^{-\lambda\Delta D} = \frac{\gamma(1+B/\beta, \lambda\Delta D)}{\Gamma(1+B/\beta)}, \tag{18}$$

which is analogous to equation (4). We can therefore design the buffer so that the probability of burst loss due to buffer overflow is less than some desired threshold ρ. For instance, if $\lambda\Delta D = 10$ bursts, and we want $\rho = 10^{-6}$, we need to have a delay buffer big enough for at least 28 bursts, which corresponds to 420 KBytes if a burst size of 15 KBytes is used.

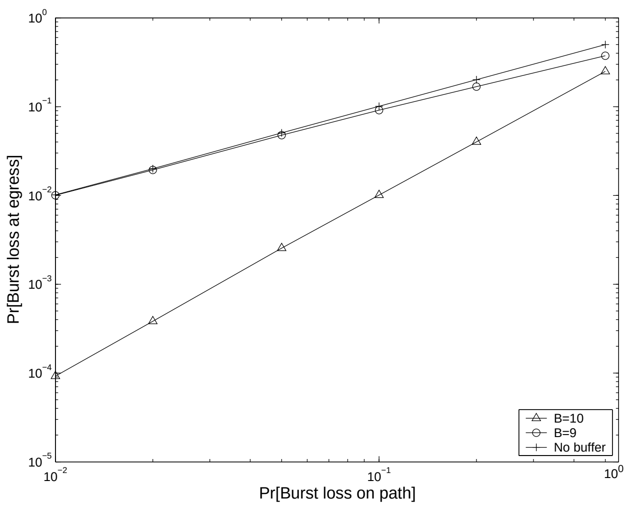

Fig. 15. Simulated burst loss probability at egress for a OBS 1+1 system with constant burst interarrival times.

We simulated the behavior of 1+1 protection scheme for OBS in which the loss probabilities on the two LSPs are assumed to be equal. Fig. 15 plots the burst loss obtained from the simulations for the case where bursts have constant interarrival times. This simulation corresponds to the timer-based burst assembly mechanism at the ingress node.

We also simulated loss behavior for the case where threshold-based burst assembly was used, leading to a burst stream that can be characterized by a Poisson process. The values obtained from that set of simulations are shown in Fig. 16.

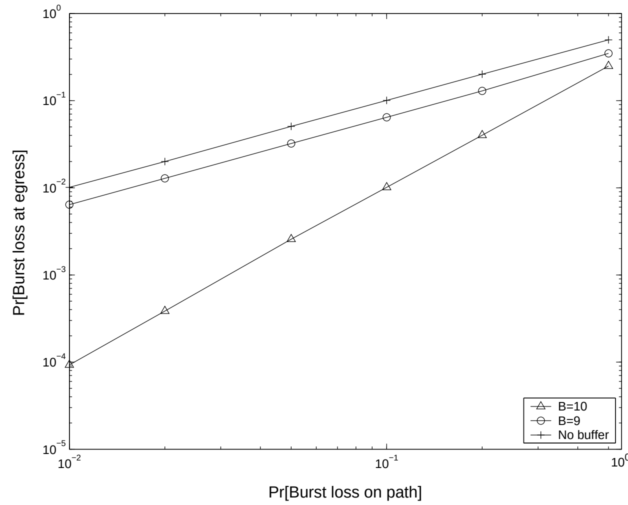

Fig. 16. Simulated burst loss probability at egress for a OBS 1+1 system with Poisson burst interarrival times.

As can be observed by the burst loss amounts for the different lengths of the buffer B in Fig. 15 and Fig. 16, the decrease in the loss is achieved by buffering the bursts at the ingress. More specifically, the burst loss for the case of $B = 10$ is lower than that for the case of $B = 1$. These results demonstrate once again the importance of taking advantage of buffering to control burst loss.

VII. Summary

In this paper we examined the MPLS 1+1 protection scheme that was introduced in [7] and used it to define a 1+1 protection architecture for OBS networks. To support the creation of protected sessions in OBS networks, some extensions are defined to the JumpStart JIT signaling architecture

On the basis of the qualitative discussion of some of the effects of delay mismatch on restoration performance in [7], we computed the gap length and number of bursts lost during the failure of the leading LSP for the most common burst assembly approaches: timer-based and threshold-based approaches. Using the deterministic and Poisson models for the two simplified burst assembly approaches, we examined the effect of delay mismatch on jitter and demonstrated that delay mismatch can introduce considerable levels of noise into the measured burst jitter, even when the jitter on the individual LSPs is small.

We also developed a model of burst loss at the OBS network egress and showed that small delay offsets can eliminate any reduction in burst loss probability that results from using duplicated bursts. The best solution to these problems appears to be using constrained routing to reduce mismatch and buffering the leading path at the ingress, rather than shortening the sliding window at the egress. Especially, benefitting from buffering at ingress was verified by simulation results. Therefore, the developed analytic model can be useful for protecting high-priority traffic in an OBS network in a manner which can satisfy the burst loss requirement. Now, we are investigating the more generalized traffic arrival distribution than Poisson and deterministic models, for burst assembly approach which applies timer-based and threshold-based methods together. As part of our future work, we plan to extend our developed analytical model in this paper to the other traffic distributions we are studying.

ACKNOWLEDGEMENTS

The authors are grateful to K. Sriram of NIST for helpful comments and suggestions, L. Krivulina of UC Berkeley for developing some of the simulations used in this study, and the anonymous reviewers for their useful comments.

APPENDIX

A. Computing $\mathsf{E}\{J\}$ with Deterministic Arrivals

In the case of random burst losses in the deterministic model, using (7), $\mathsf{E}\{|D_{n-1,n}|\,|\mathrm{same}\} = 0$ and so the expected delay variation between bursts is

$$\mathsf{E}\{|D_{n-1,n}|\} = \Delta D \cdot \mathsf{Pr}\{\mathrm{diff}\}, \tag{19}$$

where ΔD is the propagation delay difference. To compute $\mathsf{Pr}\{\mathrm{diff}\}$, we condition on which LSP produced the $(n-1)^{\mathrm{th}}$ burst forwarded by the egress LSP, giving

$$\mathsf{Pr}\{\mathrm{diff}\} = p_{n-1,n}(0,1)p_{n-1}(0) + p_{n-1,n}(1,0)p_{n-1}(1), \tag{20}$$

where $p_{m,n}(i,j)$ is the conditional probability that the n^{th} burst forwarded by the egress LSR came from LSP$_j$ given that the m^{th} burst came from LSP$_i$, and $p_n(i)$ is the probability that the n^{th} burst forwarded by the egress LSR came from LSP$_i$.

We assume bursts are dropped independently on LSP$_i$ with probability p_i. Because $D_0 < D_1$, every burst that appears on LSP$_0$ is forwarded by the egress; thus $p_{n-1}(0) = 1 - p_0$. Suppose that the $(n-1)^{\mathrm{th}}$ burst forwarded by the egress LSR is the m^{th} burst transmitted from the ingress LSR over LSP$_0$. This burst arrived at the egress LSR at time $t = D_0 + m/\lambda$, where λ is the average burst arrival rate. The n^{th} burst transmitted downstream by the egress LSR arrived at the egress from LSP$_1$ at time $t = D_1 + (m+\ell)/\lambda$, where $\ell = 1, 2, \ldots$. From (2), it follows that if no bursts are lost, burst $k+m+\ell$ will arrive at the egress from LSP$_0$ before burst $m+\ell$ arrives on LSP$_1$, where $k = \lfloor \lambda\Delta D \rfloor$ is the integer number of bursts by which LSP$_0$ is ahead of LSP$_1$, from (2).

So for burst $m+\ell$ to be selected by the egress from LSP$_0$, that burst must not be lost, while bursts $m+1, m+2, \ldots, m+k+\ell$ must be lost on LSP$_0$ and bursts $m+1, m+2, \ldots, m+\ell-1$ must be lost on LSP$_1$. The probability of this event is $(1-p_1)p_0^{k+\ell}p_1^{\ell-1}$. The total probability $p_{n-1,n}(0,1)$ is therefore

$$\begin{aligned} p_{n-1,n}(0,1) &= \sum_{\ell=1}^{\infty}(1-p_1)p_0^{k+\ell}p_1^{\ell-1} \\ &= \frac{(1-p_1)p_0^{k+1}}{1-p_0p_1}. \end{aligned} \tag{21}$$

The probability that a burst with sequence number n is accepted from LSP$_1$ is the probability that burst n was lost on LSP$_0$, along with the bursts $n+1$, $n+2$, …, $n+k$. This event occurs with probability $p_{n-1}(1) = p_0^{k+1}(1-p_1)$. Given that burst n was accepted from LSP$_1$, the probability that the next accepted burst is from LSP$_0$and has sequence number $n+k+\ell$ is the probability that bursts $n+1$, $n+2$, …, $n+\ell-1$ are lost on LSP$_1$ and bursts $n+k+1$, $n+k+2$, …, $n+k+\ell-1$ are lost on LSP$_0$, while burst $n+k+\ell$ is not lost. For a particular value of ℓ, this event occurs with probability $(1-p_0)(p_0p_1)^{\ell-1}$. If $\ell > L-k$, burst $n+k+\ell$ from LSP$_0$ will fall outside the range of the sliding window and be rejected, along with any subsequent bursts from LSP$_0$ (until the window values wrap). The conditional probability therefore is

$$\begin{aligned} p_{n-1,n}(1,0) &= \sum_{\ell=1}^{L-k}(1-p_0)(p_0p_1)^{\ell-1} \\ &= \frac{(1-p_0)(1-(p_0p_1)^{L-k})}{1-p_0p_1}. \end{aligned} \tag{22}$$

If $L \gg k$, this can be approximated as $(1-p_0)/(1-p_0p_1)$. Thus, we have

$$\mathsf{E}\{J\} = \mathsf{E}\{|D_{n-1,n}|\} = \frac{2(1-p_1)(1-p_0)}{1-p_0p_1}\Delta D p_0^{\lceil \lambda\Delta D\rceil}. \tag{23}$$

B. Computing $\mathsf{Pr}\{\mathrm{same}\}$ with Poisson Arrivals

We begin with $\mathsf{Pr}\{\mathrm{same}\}$, which can be expanded as follows:

$$\mathsf{Pr}\{\mathrm{same}\} = \mathsf{Pr}\{\text{both from LSP}_0\} + \mathsf{Pr}\{\text{both from LSP}_1\}. \tag{24}$$

We first consider $\mathsf{Pr}\{\text{both from LSP}_0\}$, which is the probability that two consecutive bursts flowing downstream from the egress traveled over LSP$_0$. We can express this as

$$\mathsf{Pr}\{\text{both from LSP}_0\} = \sum_{\ell=1}^{\infty}\mathsf{Pr}\{m,\, m+\ell \text{ from LSP}_0\}, \tag{25}$$

where m is an arbitrary integer. It is possible for us to have a situation where a burst is received from LSP$_0$ and forwarded to the destination, after which several consecutive bursts are lost on LSP$_0$ before another burst is received from that LSP and forwarded. The arrival times of bursts m and $m+\ell$ on LSP$_0$ and on LSP$_1$ are depicted in Fig. 17. The figure shows the arrival times of the two bursts on LSP$_0$ for the case where burst m arrives on LSP$_1$ before burst $m+\ell$ arrives on LSP$_0$.

To evaluate $\mathsf{Pr}\{m,\, m+\ell \text{ from LSP}_0\}$, we need to condition on the value of the time interval between the arrivals of bursts m and $m+\ell$ on LSP$_0$, so that we have

$$\begin{aligned} &\mathsf{Pr}\{m,\, m+\ell \text{ from LSP}_0\} \\ &\quad= \mathsf{Pr}\{m,\, m+\ell \text{ from LSP}_0 | Y_\ell \le \Delta D\}\,\mathsf{Pr}\{Y_\ell \le \Delta D\} \\ &\qquad+ \mathsf{Pr}\{m,\, m+\ell \text{ from LSP}_0 | Y_\ell > \Delta D\}\,\mathsf{Pr}\{Y_\ell > \Delta D\}, \end{aligned} \tag{26}$$

where Y_ℓ is defined as

$$Y_\ell = a_{m+\ell} - a_m = \sum_{i=1}^{\ell} X_{m+i}, \tag{27}$$

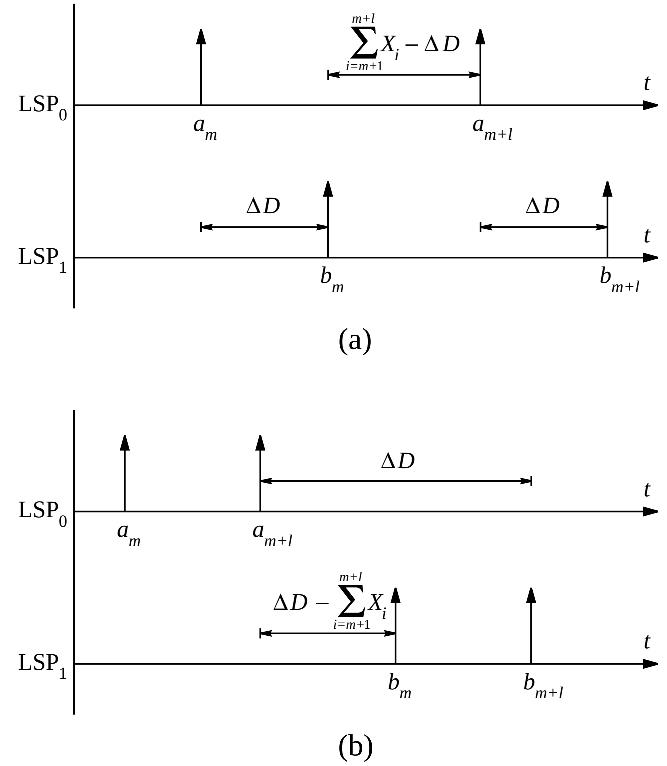

Fig. 17. Arrivals of bursts m and $m+\ell$ on LSP$_0$ and LSP$_1$ when: **(a)** $Y_\ell > \Delta D$, **(b)** $Y_\ell \leq \Delta D$.

where X_m is the amount of time that elapses between the arrivals of bursts $m-1$ and m on either LSP. (See Fig. 10.) Because Y_ℓ is the sum of ℓ independent, identically distributed exponential random variables, it has an Engset distribution of order ℓ with parameter λ. The probability that $Y_\ell \leq \Delta D$ is given by Y_ℓ's cumulative distribution function:

$$\begin{aligned}\Pr\{Y_\ell \leq \Delta D\} &= \Pr\{b_m \geq a_{m+\ell}\} \\ &= 1 - \frac{\Gamma(\ell, \lambda\Delta D)}{\Gamma(\ell)} \\ &= \frac{\gamma(\ell, \lambda\Delta D)}{\Gamma(\ell)}, \end{aligned} \tag{28}$$

and similarly

$$\Pr\{Y_\ell > \Delta D\} = \Pr\{b_m < a_{m+\ell}\} = \frac{\Gamma(\ell, \lambda\Delta D)}{\Gamma(\ell)}. \tag{29}$$

We now address each of the two cases. In both cases, the first burst arrives at time a_m and becomes the $(n-1)^{\text{th}}$ burst forwarded to the destination. If $a_{m+\ell} < b_m$, then burst $m+\ell$ is accepted and becomes the n^{th} forwarded burst if bursts $m+1, m+2, \ldots, m+\ell-1$ were dropped on LSP$_0$. Any arrivals on LSP$_1$ that occur during this interval will have no effect because their sequence numbers will fall outside the egress node's window's range, which is $[m+1, m+L]$. Thus the probability that the receiver accepts and forwards bursts m and $m+\ell$ consecutively from LSP$_0$ is

$$\Pr\{m,\, m+\ell \text{ from LSP}_0 | Y_\ell \leq \Delta D\} = (1-p_0)^2 p_0^{\ell-1}. \tag{30}$$

If $a_{m+\ell} > b_m$, as shown in Fig. 17, then burst $m+\ell$ will be forwarded from LSP$_0$ if bursts $m+1, m+2, \ldots, m+\ell-1$ are dropped on LSP$_0$ and all bursts arriving on LSP$_1$ during the time interval $[b_m, a_{m+\ell}]$ are lost. Since the sequence of arrivals on LSP$_1$ is a time-delayed copy of the arrival process on LSP$_0$, we know that the number of arrivals in the interval $[b_m, a_{m+\ell}]$ lies in the range $[0, \ell-1]$, and is the same as the number that arrive in the interval $[a_m, a_{m+\ell}]$. Furthermore, since exactly $\ell-1$ (failed) arrivals occurred in the interval $[a_m, a_{m+\ell}]$, it follows that if k arrivals occurred in the interval $[a_m, a_{m+\ell} - \Delta D]$, then $\ell-1-k$ arrivals occurred in the interval $[a_{m+\ell} - \Delta D, a_{m+\ell}]$. Thus we can compute the probability that bursts m and $m+\ell$ were received consecutively from LSP$_0$ given that $b_m < a_{m+\ell}$ as follows:

$$\begin{aligned}&\Pr\{m,\, m+\ell \text{ from LSP}_0 | Y_\ell > \Delta D\} \\ &\quad= (1-p_0)^2 p_0^{\ell-1} \sum_{k=0}^{\ell-1} p_1^k \frac{[\lambda\Delta D]^{\ell-1-k}}{(\ell-1-k)!} e^{-\lambda\Delta D} \\ &\quad= (1-p_0)^2 p_0^{\ell-1} p_1^{\ell-1} e^{-\lambda\Delta D} \sum_{m=0}^{\ell-1} p_1^{-m} \frac{[\lambda\Delta D]^m}{m!} \\ &\quad= (1-p_0)^2 p_0^{\ell-1} p_1^{\ell-1} \frac{\Gamma(\ell, \lambda\Delta D/p_1)}{(\ell-1)!}. \end{aligned} \tag{31}$$

We are now in a position to evaluate (25), which we do in two parts. Combining the results obtained in (28) and (30) and summing over all values of ℓ, we get

$$\begin{aligned}&\sum_{\ell=1}^{\infty} \frac{\gamma(\ell, \lambda\Delta D)}{(\ell-1)!}(1-p_0)^2 p_0^{\ell-1} \\ &\quad= (1-p_0)^2 \int_0^{\lambda\Delta D} e^{-u} \left[\sum_{\ell=1}^{\infty} \frac{[p_0 u]^{\ell-1}}{(\ell-1)!}\right] du \\ &\quad= (1-p_0)\left(1 - e^{-\lambda(1-p_0)\Delta D}\right), \end{aligned} \tag{32}$$

since the summation in the integrand reduces to $e^{p_0 u}$. We perform a similar operation by combining (29) and (31) and summing over all values of ℓ, which gives us

$$\begin{aligned}&(1-p_0)^2 \sum_{\ell=1}^{\infty} \frac{\Gamma(\ell, \lambda\Delta D)}{(\ell-1)!} \frac{\Gamma(\ell, \lambda\Delta D/p_1)}{(\ell-1)!} p_0^{\ell-1} p_1^{\ell} \\ &\quad= (1-p_0)^2 \int_{\lambda\Delta D}^{\infty} \int_{\lambda\Delta D/p_1}^{\infty} e^{-(t+u)} I_0(2\sqrt{p_0 p_1 t u})\, dt du \end{aligned} \tag{33}$$

where $I_0(z)$ is the modified Bessel function of the first kind and order 0, given by

$$I_0(z) = \sum_{k=0}^{\infty} \frac{(z^2/4)^k}{(k!)^2}. \tag{34}$$

For $\Pr\{\text{both from LSP}_1\}$, we assume that the $(n-1)^{\text{th}}$ burst forwarded by the egress was the burst with sequence number m, which this time was correctly received from LSP$_1$. The n^{th} burst forwarded by the egress will be burst $m+\ell$, also received from LSP$_1$, where ℓ is a positive integer. For this to happen, bursts $m, m+1, \ldots, m+\ell$ must be lost on LSP$_0$, along with all bursts arriving on LSP$_0$in the time interval $[a_{m+\ell}, b_{m+\ell}]$, and bursts $m+1, m+2, \ldots, m+\ell-1$ must be lost on LSP_1. Thus the probability of bursts with sequence numbers m and $m+\ell$ being the $(n-1)^{\mathrm{th}}$ and n^{th} bursts forwarded to the destination is

$$\begin{aligned}
&\mathsf{Pr}\left\{m,m+\ell \text{ from } \mathrm{LSP}_1\right\}\\
&= (1-p_1)^2 p_1^{\ell-1} p_0^{\ell+1} \sum_{k=0}^{\infty} p_0^k \frac{[\lambda\Delta D]^k}{k!} e^{-\lambda\Delta D}\\
&= (1-p_1)^2 p_1^{\ell-1} p_0^{\ell+1} e^{-\lambda(1-p_0)\Delta D}. \qquad (35)
\end{aligned}$$

Summing over all values of ℓ gives us

$$\mathsf{Pr}\left\{\text{both from } \mathrm{LSP}_1\right\} = \frac{p_0^2(1-p_1)^2}{1-p_0p_1} e^{-\lambda(1-p_0)\Delta D}. \qquad (36)$$

We finally get $\mathsf{Pr}\{\text{same}\}$ by summing (32), (33), and (36).

C. *Computing* $\mathsf{Pr}\{\text{diff}\}$ *with Poisson Arrivals*

Next we compute $\mathsf{Pr}\{\text{diff}\}$, which is the probability that two subsequent bursts received by the destination came from different protection LSPs. This can be expanded as

$$\begin{aligned}
\mathsf{Pr}\{\text{diff}\} &= \mathsf{Pr}\left\{\text{first from } \mathrm{LSP}_0, \text{ second from } \mathrm{LSP}_1\right\}\\
&\quad + \mathsf{Pr}\left\{\text{first from } \mathrm{LSP}_1, \text{ second from } \mathrm{LSP}_0\right\}. \qquad (37)
\end{aligned}$$

First consider $\mathsf{Pr}\left\{\text{first from } \mathrm{LSP}_0, \text{ second from } \mathrm{LSP}_1\right\}$, which is

$$\begin{aligned}
&\mathsf{Pr}\left\{\text{first from } \mathrm{LSP}_0, \text{ second from } \mathrm{LSP}_1\right\}\\
&= \sum_{\ell=1}^{\infty} \mathsf{Pr}\left\{m \text{ from } \mathrm{LSP}_0,\ m+\ell \text{ from } \mathrm{LSP}_1\right\}, \qquad (38)
\end{aligned}$$

for some integer m. In order to determine $\mathsf{Pr}\left\{m \text{ from } \mathrm{LSP}_0, m+\ell \text{ from } \mathrm{LSP}_1\right\}$, we note that if burst m is successfully received from LSP_0 with probability $1-p_0$, what happens to the copy of burst m that arrives from LSP_1 does not matter because the egress' window will have moved so that the sequence number m is no longer in range. For the next burst that is forwarded by the egress LSR to be burst $m+\ell$ from LSP_1, bursts $m+1, m+2, \ldots, m+\ell$ must be dropped on LSP_0, along with all the bursts that arrive on LSP_0 in the time interval $[a_{m+\ell}, b_{m+\ell}]$, which has length ΔD. On LSP_1, bursts $m+1, m+2, \ldots, m+\ell-1$ must be lost also. From these criteria, we can write the probability that the first of consecutive forwarded bursts is from LSP_0 and the second is from LSP_1as

$$\begin{aligned}
&\sum_{\ell=1}^{\infty} \mathsf{Pr}\left\{m \text{ from } \mathrm{LSP}_0,\ m+\ell \text{ from } \mathrm{LSP}_1\right\}\\
&= \sum_{\ell=1}^{\infty} (1-p_0) p_0^{\ell} \left[\sum_{k=0}^{\infty} p_0^k \frac{[\lambda\Delta D]^k}{k!} e^{-\lambda\Delta D}\right] p_1^{\ell-1}(1-p_1)\\
&= \frac{(1-p_0)(1-p_1)p_0 e^{-\lambda(1-p_0)\Delta D}}{1-p_0p_1}. \qquad (39)
\end{aligned}$$

Note that in this case there is no dependence on the value taken by $Y_\ell = a_{m+\ell} - a_m$.

In a similar fashion we can compute $\mathsf{Pr}\left\{\text{first from } \mathrm{LSP}_1, \text{ second from } \mathrm{LSP}_0\right\}$. In this case, we must condition on the value taken by Y_ℓ, because this event can occur only if $Y_\ell > \Delta D$. The probability is

$$\begin{aligned}
&\mathsf{Pr}\left\{\text{first from } \mathrm{LSP}_1, \text{ second from } \mathrm{LSP}_0\right\}\\
&= \sum_{\ell=1}^{\infty} \mathsf{Pr}\left\{m \text{ from } \mathrm{LSP}_1,\ m+\ell \text{ from } \mathrm{LSP}_0 | Y_\ell > \Delta D\right\}\\
&\qquad \times \mathsf{Pr}\left\{Y_\ell > \Delta D\right\}, \qquad (40)
\end{aligned}$$

where $\mathsf{Pr}\{Y_\ell > \Delta D\}$ is given by (29). Burst m will be forwarded from LSP_1 if its copy is not successfully received from LSP_0 and all bursts arriving on LSP_0 in the time interval $[a_m.b_m]$ are also lost. Also, burst $m+\ell$ will be forwarded from LSP_0 if it is received successfully with probability $1-p_1$, all $\ell-1$ bursts arriving on LSP_0 in the interval $(a_m, a_{m+\ell})$ are lost, and all bursts arriving on LSP_1 in the interval $(b_m, a_{m+\ell})$ are lost as well. Because there are $\ell-1$ arrivals in the interval $(b_m, b_{m+\ell})$, we know that if k arrivals occur in $(b_m, a_{m+\ell})$, then there must be $\ell-1-k$ arrivals in the interval $(a_{m+\ell}, b_{m+\ell})$, where $0 \le k \le \ell-1$. Thus we get

$$\begin{aligned}
&\mathsf{Pr}\left\{m \text{ from } \mathrm{LSP}_1,\ m+\ell \text{ from } \mathrm{LSP}_0 | Y_\ell > \Delta D\right\}\\
&= (1-p_0)(1-p_1)p_0^{\ell} \sum_{k=0}^{\ell-1} p_1^k \frac{[\lambda\Delta D]^{\ell-1-k}}{(\ell-1-k)!} e^{-\lambda\Delta D}\\
&= (1-p_0)(1-p_1)p_0^{\ell} p_1^{\ell-1} e^{-\lambda(1-p_1^{-1})\Delta D} \frac{\Gamma(\ell, \lambda\Delta D/p_1)}{(\ell-1)!}. \qquad (41)
\end{aligned}$$

Summing over all values of ℓ gives us

$$\begin{aligned}
&\mathsf{Pr}\left\{\text{first from } \mathrm{LSP}_1, \text{ second from } \mathrm{LSP}_0\right\}\\
&= (1-p_0)(1-p_1)e^{-\lambda(1-p_1^{-1})\Delta D}\\
&\quad \times \sum_{\ell=1}^{\infty} p_0^{\ell} p_1^{\ell-1} \frac{\Gamma(\ell, \lambda\Delta D/p_1)}{(\ell-1)!} \frac{\Gamma(\ell, \lambda\Delta D)}{(\ell-1)!}\\
&= (1-p_0)(1-p_1)p_0 e^{-\lambda(1-p_1^{-1})\Delta D}\\
&\quad \times \int_{\lambda\Delta D/p_1}^{\infty} \int_{\lambda\Delta D}^{\infty} e^{-(t+u)} I_0(2\sqrt{p_0p_1tu})\, dt du \qquad (42)
\end{aligned}$$

We obtain $\mathsf{Pr}\{\text{diff}\}$ by combining (39) and (42).

References

[1] A. Banerjee, J. Drake, J.P. Lang, B. Turner, K. Kompella, and Y. Rekhter, "Generalized multiprotocol label switching: An overview of routing and management enhancements," *IEEE Communications Magazine*, vol. 39, no. 1, pp. 144–150, January 2001.

[2] V. Sharma and F. Hellstrand (Eds.), "Framework for MPLS-based recovery," IETF RFC 3469, February 2003.

[3] E. Mannie et al., "Recovery (protection and restoration) terminology for GMPLS," IETF Internet Draft, 2002.

[4] C. Huang, V. Sharma, K. Owens, and S. Makam, "Building reliable MPLS networks using a path protection mechanism," *IEEE Communications Magazine*, vol. 40, no. 3, pp. 156 – 162, 2002.

[5] R. Bartos and M. Raman, "A heuristic approach to service restoration in MPLS networks," in *Proceedings of the IEEE International Conference on Communications 2001 (ICC 2001)*, 2001, pp. 117–121.

[6] D. Zhou and S. Subramaniam, "Survivability in optical networks," *IEEE Network*, vol. 14, no. 6, pp. 16–23, Nov–Dec 2000.

[7] "Architecture and specification of data communication network," ITU-T Prepublished Recommendation G.7712/Y.1703, March 2003.

[8] N. F. Maxemchuk, "Dispersity routing," in *Proceedings of the International Conference on Communications 1975 (ICC '75)*, 1975, pp. 41–10–41–13.

[9] Y.A. Kim, J.K. Choi, and S.G. Jong, "Analysis of the end-to-end recovery algorithm in the IP over WDM network," in *Proceedings of the International Conference on Optical Internet (COIN)*, 2002.
[10] E. Kozlovski and P. Bayvel, "Link failure restoration in wavelength-routed optical burst switched (wr-obs) networks," in *Optical Fiber Communications (OFC 2003)*, 2003.
[11] S. Oh, Y. Kim, M. Yoo, and H. H. Hong, "Survivability in the optical internet using the optical burst switch," *ETRI Journal*, vol. 24, no. 2, pp. 117–130, 2002.
[12] F. Callegati A. Ge and L. S. Tamil, "On optical burst switching and self-similar traffic," vol. 4, no. 3, pp. 98–100, 2000.
[13] K. Haridoss V. M. Vokkarane and J. P. Jue, "Threshold-based burst assembly policies for qos support in optical burst-switched networks," in *Proceedings of the SPIE Optical Communications Conference (Opti-Comm 2002)*, 2002, vol. 4874, pp. 125–136.
[14] D. Griffith and S. Lee, "1+1 protection for optical burst switched networks," in *submitted to Globecom 2003*.
[15] R. Bhandari, "Optimal physical diversity algorithms and survivable networks," in *Proceedings of the Second IEEE Symposium on Computers and Communications, 1997*, 1997, pp. 433–441.
[16] C. Qiao, "Labeled optical burst switching for IP-over-WDM integration," *IEEE Communications Magazine*, vol. 38, no. 9, pp. 104 – 114, 2000.
[17] M. Yoo, C. Qiao, and S. Dixit, "Optical burst switching for service differentiation in the next-generation optical internet," *IEEE Communications Magazine*, vol. 39, no. 2, pp. 98 – 104, 2001.
[18] I. Baldine, G.N. Rouskas, H.G. Perros, and D. Stevenson, "JumpStart: a just-in-time signaling architecture for WDM burst-switched networks," *IEEE Communications Magazine*, vol. 40, no. 2, pp. 82–89, 2002.
[19] M. Yoo, C. Qiao, and S. Dixit, "QoS performance of optical burst switching in IP-Over-WDM networks," *IEEE Journal on Selected Areas in Communications*, vol. 18, no. 10, pp. 2062 – 2071, 2000.
[20] "JumpStart JIT signaling definition, revision 1.17," Tech. Rep., MCNC-RDI ANR and North Carolina State University, May 2003, URL: http://jumpstart.anr.mcnc.org/Docs/03_05_12/SignalingDefinition.pdf.
[21] Z. Rosberg, H. L. Vu, M. Zukerman, and J. White, "Blocking probabilities of optical burst switching networks based on reduced load fixed point approximations," in *Proceedings of the Twenty-Second Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM 2003)*, 2003.
[22] H. Schulzrinne, S. Casner, R. Frederick, and V. Jacobson, "RTP: A transport protocol for real-time applications," IETF RFC 1889, January 1996.